

## FORMATION OF THE BLACK HOLES IN THE HIGHEST REDSHIFT QUASARS

JAIYUL YOO AND JORDI MIRALDA-ESCUDÉ

Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210;  
 jaiyul, jordi@astronomy.ohio-state.edu

*accepted for publication in The Astrophysical Journal Letters*

### ABSTRACT

The recent discovery of luminous quasars up to a redshift  $z = 6.43$  has renewed interest in the formation of black holes massive enough to power quasars. If black holes grow by Eddington-limited gas accretion with a radiative efficiency of at least 10%, the time required to grow from a stellar black hole to  $\sim 10^9 M_\odot$  is  $\sim 10^9$  years, close to the age of the universe at  $z = 6$ . Black hole mergers may accelerate the rate of mass growth, but can also completely eject black holes from halo centers owing to the gravitational wave recoil effect. Recently, Haiman concluded that black hole ejections likely do not allow black holes to grow to  $\sim 10^9 M_\odot$  by  $z = 6.43$ . We reexamine this problem and show that, by using a different halo escape velocity, accounting for the dependence of the recoil velocity on the black hole binary mass ratio and spins, and allowing seed black holes to form in all halos down to virial temperatures of 2000 K, black hole masses may reach  $\sim 10^9 M_\odot$  as early as  $z = 9$  starting from stellar seeds, without super-Eddington accretion. In this particular case, we find that these massive black holes form from the merger of  $\sim 10^4$  stellar black holes formed in low-mass halos at  $z \sim 20$ , which must all grow close to the maximum Eddington rate over most of the time available from their birth to  $z \sim 6$ . The alternative is that black holes can grow more rapidly by super-Eddington accretion.

*Subject headings:* black hole physics — cosmology: theory — galaxies: nuclei — gravitation — gravitational waves

### 1. INTRODUCTION

Highly luminous quasars at high redshift require the formation of black holes (BH) with a mass  $M$  large enough to make the observed luminosity be less than the Eddington limit. This may pose a problem for cold dark matter (CDM) models where massive halos that can harbor BHs do not form until late epochs (Efstathiou & Rees 1988). Currently, the highest redshift quasar known is SDSS 1148+3251, at  $z = 6.43$  (Fan et al. 2003). Its luminosity  $L$ , if equated to the Eddington value  $L_{\text{Edd}} = 4 \times 10^4 L_\odot (M/M_\odot)$ , results in a minimum BH mass  $M = 4 \times 10^9 M_\odot$  (Fan et al. 2003; Haiman 2004). The quasar luminosity inferred from the observed flux is not likely to be greatly overestimated due to gravitational lensing or beaming (Haiman & Cen 2002; Willott et al. 2003; Keeton et al. 2004).

If BHs grew in mass by accretion through a standard thin disk with radiative efficiency  $\epsilon \sim 0.1$ , then a BH of initial mass  $M_0$  formed at time  $t_0$  accreting continuously up to time  $t$  with luminosity  $L$  will reach a final mass  $M = M_0 \exp[(t-t_0)/t_{\text{Sal}}]$ , where  $t_{\text{Sal}}$  is the Salpeter time (Salpeter 1964),

$$t_{\text{Sal}} = \frac{\epsilon M c^2}{(1-\epsilon)L} = 4 \times 10^7 \text{ yr} \frac{\epsilon}{0.1(1-\epsilon)} \frac{L_{\text{Edd}}}{L}. \quad (1)$$

Note that the factor  $1-\epsilon$  in the denominator is to account for the fact that only a fraction  $1-\epsilon$  of the accreted mass is added to the BH. Since the age of the universe at  $z=6$  is  $\sim 10^9$  years, a BH made by one of the first stars formed in the universe with  $M_0 \sim 10 M_\odot$  would need to be accreting near the Eddington rate for almost all the time available before  $z=6$  in order to increase its mass by a factor  $\sim 10^8$  to power the observed quasars (corresponding to  $\sim 20$  e-folding times).

There are several ways to ease this requirement and allow BHs to grow to a large mass in less time: (a) BHs may have started at a higher initial mass from the collapse of supermassive stars (e.g., Carr, Bond & Arnett 1984). (b) There may be super-Eddington accretion with  $t_{\text{Sal}} < 4 \times 10^7$  years if the

radiative efficiency is  $\epsilon < 0.1$  (e.g., Ohsuga et al. 2002), or the luminosity is greater than the Eddington luminosity with  $\epsilon = 0.1$  (Ruszkowski & Begelman 2003). (c) Finally, several BHs may have merged to form a more massive BH.

In this paper we consider the latter possibility in the context of the CDM model, where dark matter halos merge hierarchically. If the earliest BHs arose from the first stars forming in the center of halos, then halo mergers can lead to the merger of their central BHs (e.g., Volonteri et al. 2003), after dynamical friction leads to the formation of a binary BH which can then lose orbital energy by interacting with stars (Begelman, Blandford & Rees 1980) or via gas dissipation (Gould & Rix 2000), until the emission of gravitational waves takes over. If many BHs growing from initial stellar seeds of mass  $M_0$ , made at time  $t_0$ , continuously accrete until time  $t$  (as they merge with each other at any intermediate times), the final BH can reach a mass (Haiman & Loeb 2001),

$$M = \sum_i M_{0,i} e^{(t-t_0)/t_{\text{Sal}}}. \quad (2)$$

In the last stage of the merger of BHs, the emission of gravitational waves (GW) can give a net momentum to the resulting BH large enough to eject it from its halo (Fitchett 1983; Kidder 1995; Favata et al. 2004; Madau et al. 2004; Merritt et al. 2004). Then the halo center is left empty and the process of BH growth must start over again. Recently, Haiman (2004) presented a calculation of the growth of BH masses assuming that they start growing from stellar seeds only when the halo has a velocity dispersion  $\sigma > v_{\text{kick}}/2$ , where  $v_{\text{kick}}$  is a fixed kick velocity given to the BHs by the final GW burst. Haiman (2004) concluded that for  $v_{\text{kick}} > 64 \text{ km s}^{-1}$ , BHs would not be able to grow to large enough masses without super-Eddington accretion.

We show in this paper that when the adequate kick velocity as a function of the mass ratio and spins of the two merging BHs is used, and a halo escape velocity based on an isothermal profile (higher than assumed by Haiman) is used, ejec-

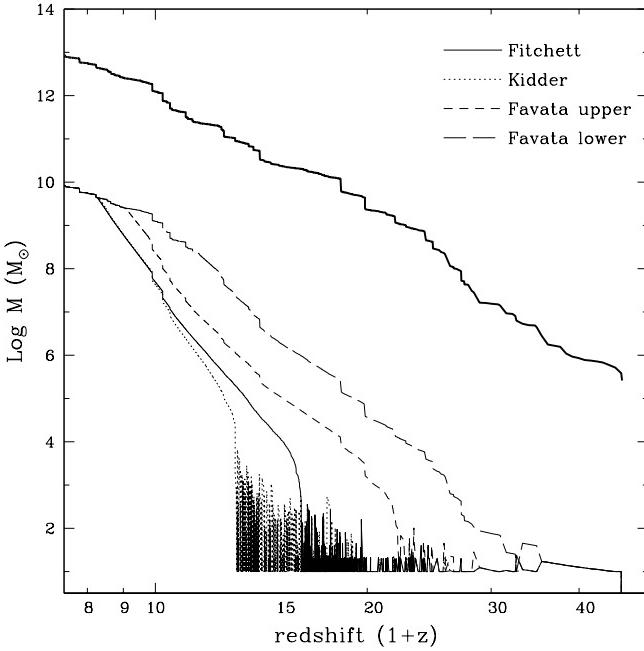


FIG. 1.— Mass of the main BH progenitor as a function of redshift ending at  $z = 6.43$ , according to four different estimates of GW recoil effect (shown as the four lower curves). The upper curve (thick solid line) shows the mass of the main halo progenitor. Seed BHs are started at  $10M_{\odot}$ , and the maximum BH mass allowed is  $10^{-3}$  of the halo mass. For reference, GW recoil velocities are of order  $1000, 1000, 400$ , and  $50 \text{ km s}^{-1}$  for Fitchett, Kidder, Favata upper and lower models, respectively.

tions by GW emission result in a smaller reduction than found by Haiman (2004) of the final black hole mass that is attained. As a result, we shall show that BH masses of  $10^9 M_{\odot}$  can be reached up to  $z \simeq 9$  without super-Eddington accretion, even for the highest plausible kick velocities found by Favata et al. (2004). Our model for the BH evolution and growth is described in § 2. The main results are presented in § 3, and we discuss the implications in § 4.

## 2. BLACK HOLE EVOLUTION MODEL

We model the evolution of dark matter halos using the merger tree of Cole et al. (2000) based on the extended Press-Schechter formalism (Press & Schechter 1974; Bond et al. 1991; Lacey & Cole 1993). We use the flat  $\Lambda$ CDM power spectrum (Eisenstein & Hu 1999) with matter density  $\Omega_m = 0.3$ , baryon density  $\Omega_b = 0.04$ , present Hubble constant  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , power spectrum normalization  $\sigma_8 = 0.9$ , and primordial index  $n = 1$ , consistent with the WMAP results (Bennett et al. 2003; Spergel et al. 2003). We start the merger tree at  $z = 6.43$  with a final halo mass  $M_f$ , finding all the progenitor halos at higher redshift. The mass  $M_f$  is determined by requiring the quasar number density  $n_Q = 2.7 \times 10^{-9} (h^{-1} \text{ Mpc})^3$  (Fan et al. 2003) to be equal to the number density of halos of mass  $M_f$  at  $z = 6.43$ . This gives the maximum possible mass of the host halo of the quasar consistent with the observed quasar abundance (for example, if quasars have a short duty cycle the host halo abundance must be higher than the quasar abundance, so the halo mass must be lower). The resulting halo mass,  $M_f \simeq 8.5 \times 10^{12} M_{\odot}$  (Haiman 2004), requires a  $5.2\sigma$  fluctuation at  $z = 6.43$ .

The first BHs are assumed to form in halos that have reached a virial temperature  $T = 2000 \text{ K}$ , when molecular hy-

drogen cooling induces the formation of the first stars (e.g., Yoshida et al. 2003). A constant initial BH mass  $M_0$  is assumed in every halo reaching this virial temperature. The mass corresponding to this virial temperature,  $M_{res}(z)$ , is chosen as the resolution of the merger tree. In order to accurately follow the appearance and growth of halos down to mass  $M_{res}$ , we compute accretion into halos of mass  $M_h$  from smaller halos down to a minimum mass  $10^{-3} M_h$  or  $M_{res}$ , whichever is smallest. We compute the virial temperature of halos, related to velocity dispersion as  $T = \mu \sigma^2 / k$  (where  $k$  is Boltzmann's constant and  $\mu$  is the mean particle mass), using the fitting formula of Bryan & Norman (1998) for the mean overdensity of virialized halos  $\Delta_c$ , and the equations  $\sigma^2 = GM_h / (2R_h)$ ,  $M_h = (4\pi/3)R_h^3 \rho_c(z) \Delta_c(z)$ , where  $R_h$  and  $\rho_c$  are the radius of virialization of the halo and the critical density at a given redshift  $z$ , respectively.

We assume that all the BHs grow continuously by gas accretion on the timescale  $t_{\text{Sal}}$  obtained with  $\epsilon/(1-\epsilon) = 0.1$  and  $L = L_{\text{Edd}}$ , after their host halo has reached a minimum velocity dispersion  $\sigma_{min}$ . We shall show some results for  $\sigma_{min} = 0$ , in which case BHs start growing as soon as they are formed. Alternatively, the effects of slow gas cooling, stellar winds, or supernovae may prevent any gas concentrating in the halo center and being captured by the BH for accretion when  $\sigma < \sigma_{min}$ . The BH growth by gas accretion continues until the BH mass reaches a fraction  $10^{-3}$  of its halo mass, at which point we simply set  $M = 10^{-3} M_h$ .

The last step is to check if the BH merger results in an ejection from the host halo. We consider four estimates of the GW recoil velocity. The first is a quasi-Newtonian calculation of non-spinning BHs by Fitchett (1983). Second, Kidder (1995) added a post-Newtonian spin-orbit correction to Fitchett's work that depends on the BH spins. Lastly, Favata et al. (2004) and Merritt et al. (2004) obtained a new estimate using

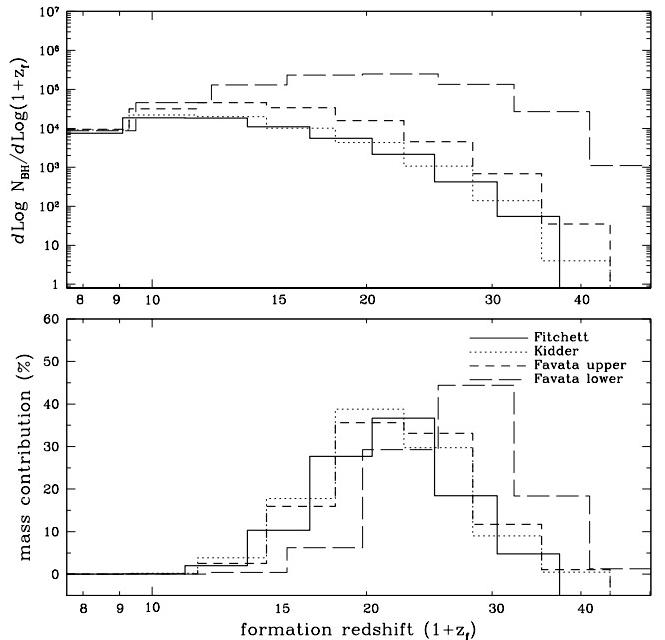


FIG. 2.— Composition of the final BH. *Upper panel:* Number of seed BHs that merged into the final BH, as a function of their formation redshift. *Lower panel:* Contribution to the final mass of the BH from the BH seeds as a function of their formation redshift.

BH perturbation theory, and they provided a lower and upper limit on the GW kick velocity comprising the uncertainty due to the final plunging state which gives the dominant contribution (see their eqs. [1], [2]). We use these limits as two separate estimates.

The mass accretion histories of BHs are calculated following the merger trees in time and computing the GW recoil velocity every time two halos merge (and by assumption, their central BHs merge). The dimensionless spin parameters of merging BHs are uniformly chosen from antiparallel ( $\hat{a} = -1$ ) to parallel ( $\hat{a} = 1$ ) in relation to the orbital angular momentum, to calculate the effective spin parameters (Damour 2001; Damour & Buonanno 2000). We then compare the GW kick velocity with the escape velocity of the halo to decide if the merged BH is ejected. We compute the escape velocity assuming the halo has an isothermal profile, with density  $\rho \propto r^{-2}$  down to the radius of the BH zone of influence:  $R_z \equiv GM/2\sigma^2$ . Therefore, the escape speed at  $R_z$  is  $v_{esc} = 2\sigma\sqrt{1 + \ln(R_h/R_z)}$ , or about  $5\sigma$  when  $M = 10^{-3}M_h$ . Note that Haiman (2004) assumed instead that  $v_{esc} = 2\sigma$ , and that no BHs start growing until the halo escape velocity has exceeded an assumed, fixed GW kick velocity. The exact value of the escape velocity depends on the halo density profile, but it would not be much lower than we assume for a profile similar to observed galaxies. Gravitational waves emission also implies that a fraction of the mass in BHs is lost as the energy of the waves. This loss of energy does not greatly reduce the final BH mass (Menou & Haiman 2004), and we neglect it in this paper.

### 3. RESULTS

Figure 1 shows the mass history of the BH in the  $z = 6.43$  quasar SDSS 1148+5251 in our models. The upper thick solid line shows the mass of the main halo progenitor in one realization of the merger tree, obtained by always choosing the branch of the most massive progenitor at every merger. The lower four lines show the mass of the BH in this main halo progenitor, according to our four different prescriptions for the GW kick velocity. The mass of the seed BHs is fixed at  $M_0 = 10M_\odot$ , and they start growing by gas accretion immediately as they are formed, over a time  $t_{Sal} = 4 \times 10^7$  years. At every halo merger, the mass of the two BHs is added when the GW kick velocity is smaller than the escape velocity. This results in the sudden BH mass increases seen in the figure. Growth by gas accretion is assumed to continue immediately after mergers. When the GW kick velocity exceeds the escape velocity, the BH is removed and replaced immediately with a new seed with  $M_0 = 10M_\odot$ . The figure shows that BHs are often kicked out at high redshift when they reside in low-mass halos, but this stops when the halo reaches a high enough escape velocity. Large, sudden mass increases (for example, at  $z \simeq 13$  for the Kidder curve) are due to additions of BHs that have grown to high mass in a merging subhalo. Note that the Favata lower limit curve reaches a mass  $10^{-3}M_h$  (at which point the BH is not allowed to grow further) at  $z \simeq 10$ .

Figure 2 shows a histogram of the number of BHs that have merged into the final BH at  $z = 6.43$ , as a function of their formation redshift (*top panel*), and their contribution to the total mass of the final BH (*bottom panel*). In other words, the number shown in the top panel is the number of BHs that we need to add over in equation (2) from every interval of BH formation redshift, and the bottom panel shows their contribution to the final mass including the factor  $e^{(t-t_{0i})/t_{Sal}}$ . The figure shows that the final mass attained by the BH at  $z = 6.43$  is the result

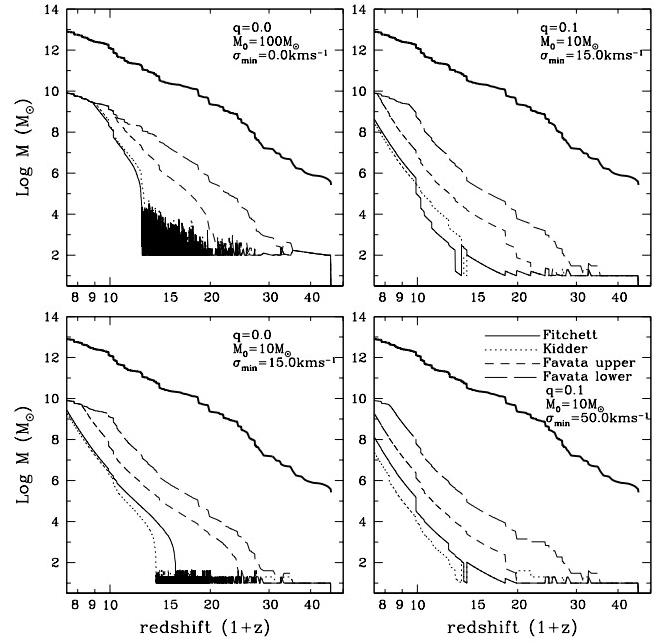


FIG. 3.— BH mass histories as in Fig. 1, changing model parameters for mergers and growth by gas accretion. *Upper left panel:* the seed BH mass is increased from  $10M_\odot$  to  $100M_\odot$ . *Lower left panel:* Growth by gas accretion starts only once the host halo has reached a minimum velocity dispersion  $\sigma_{min} = 15 \text{ km s}^{-1}$ . *Upper right panel:* BH mergers occur only when merging halos have a mass ratio higher than  $q = 0.1$ . *Lower right panel:* The minimum velocity dispersion is increased to  $\sigma_{min} = 50 \text{ km s}^{-1}$ .

of the merger of many small BHs: the formation redshift bins with the highest contribution to the final BH mass typically contain  $10^4$  to  $10^5$  seed black holes that merged. Hence, the  $\sim 10^8$  mass growth factor from the initial seed mass of  $10M_\odot$  to the final BH mass of  $\sim 10^9M_\odot$  originates in about equal parts from the growth factor  $e^{(t-t_{0i})/t_{Sal}}$  of each individual BH, and from the number of similar terms that are added together in equation (2) from all the BHs that have merged.

In summary, these results show that in the optimistic model we have assumed for BH growth, it is possible to reach a mass as high as that inferred for the BH in the quasar SDSS 1148+5251 at a redshift of up to  $z \simeq 10$  for the Favata et al. (2004) lower limit of the GW kick velocity. Even for the highest GW kick velocities the same BH mass can still be reached at  $z \simeq 8$ .

In Figure 3 we examine a number of variations on our basic model. Increasing the seed BH mass to  $100M_\odot$  (*upper left panel*) causes a small increase of the redshift at which a mass of  $\sim 10^9M_\odot$  can be achieved. The models in the lower left panel show what happens when growth by gas accretion is allowed only in halos with  $\sigma_{min} > 15 \text{ km s}^{-1}$ , since it may seem implausible that BHs accrete at the Eddington rate in halos in which the gas can be easily blown out after being photoionized. In this case, the Fitchett and Kidder models for the GW kick velocities can reach a maximum mass of only  $\sim 2 \times 10^9M_\odot$  by  $z = 6.43$ . In the upper right panel, we impose an additional requirement: halo mergers will lead to the mergers of their central BHs only if the mass ratio of the two halos,  $q$ , is between 0.1 and 1. Otherwise the merging halo is assumed to remain an orbiting satellite (with too long a dynamical friction timescale for spiraling in), so its BH never

reaches the center. In this case the Fitchett and Kidder models reach a mass of only  $3 \times 10^8 M_\odot$  at  $z = 6.43$ , and with the upper limit to GW kick velocities of Favata et al. the required mass of  $4 \times 10^9 M_\odot$  is barely reached, but for the Favata et al. lower limit to the GW kick velocities the required mass can still be reached at  $z \simeq 9$ .

We also show the result of raising  $\sigma_{min}$  to 50 km s $^{-1}$  in the lower right panel, and maintaining the requirement  $q > 0.1$  on the halo mass ratio. This higher limit might be more realistic if central BHs are only fed in galaxies with gravitationally unstable disks and a sufficient star formation rate (so that the gas does not all settle in a large, quiescent disk), and where the gravitational potential is sufficiently deep. Even in this case, the Favata et al. lower limit to the kick velocity can produce a BH mass above  $4 \times 10^9 M_\odot$  by  $z = 6.43$ , but with higher kick velocities the BH mass that can be reached is too low. Note that in the models of Haiman (2004) BHs do not appear at all when  $\sigma < \sigma_{min}$ , whereas in our model BHs can still form and merge (but not accrete gas) for  $\sigma < \sigma_{min}$ . This, together with the higher  $v_{esc}$  we adopt, explains the difference between Haiman's results and ours.

#### 4. CONCLUSIONS

Combining a continued mass growth of BHs from Eddington-limited gas accretion at high radiative efficiency ( $\epsilon \simeq 0.1$ ) with BH mergers, we can still account for the presence of BHs with  $\sim 10^9 M_\odot$  at  $z = 6.43$ , starting from stellar black holes of  $10 M_\odot$ . Typically, these BHs can form by the merger of  $\sim 10^4$  stellar BH seeds, growing by a factor  $\sim 10^4$  in mass by gas accretion. We have shown that the ejection of some BHs due to GW recoil does not necessarily impede the growth of BHs in halo centers.

Our model for the formation of these BHs involves two optimistic assumptions: First, that the BH masses grow by accretion with an e-folding time shorter than  $t_{Sal} \sim 4 \times 10^7$  years, for most of the time from their birth as stellar black holes at  $z \sim 20$  to  $z \simeq 6$ . Second, that BH mergers in which the recoil velocity is less than the halo escape velocity result in the BH immediately returning to the center and continuing its gas accretion, without significant losses from the en-

ergy of the emitted GWs. These two assumptions are highly questionable. The first one may seem unlikely in view of the small fraction of galaxies that host quasars at lower redshift, although this fraction might be high in the earliest massive halos to form. Regarding the second, a more realistic calculation might show that black holes spend a lot of time finding their way back to the halo center after the GW kick, even when they do not escape the halo (Hut & Rees 1992; Merritt et al. 2004; Madau & Quataert 2004).

The earliest redshift at which a mass of  $\sim 10^9 M_\odot$  can be reached strongly depends on the value of  $\epsilon$  assumed. Making  $\epsilon$  slightly greater than 0.1 would increase  $t_{Sal}$  and reduce the number of e-folding times available between the epoch of formation of the first stellar black holes at  $z \sim 20$  and the earliest time when luminous quasars are observed. This would then make it impossible to attain the required mass. At the same time, that is perfectly possible if black holes gain their mass in a short period of super-Eddington accretion, which is in fact the main alternative to the model of continuous Eddington-rate growth we have assumed here.

In the absence of super-Eddington accretion, the scenario of continuous, luminous gas accretion and large numbers of mergers implies that quasars should be strongly correlated at high redshift. The halo hosting the luminous quasar at  $z = 6.43$  should have merged from lower mass halos in the recent past. For example, in the randomly generated merging history used for the realization shown in Figure 1, the halo mass at  $z = 9$  is only about one third of its mass at  $z = 6.43$ , and several halos of mass greater than  $10^{11} M_\odot$  merge with the main halo in the intervening time. Most of these halos would need to host their own quasar shining near the Eddington luminosity for most of the time, within a distance of the turn-around radius of the halo at  $z = 6.43$  (a few comoving megaparsecs).

We are grateful to Zheng Zheng and Adam Steed for discussions. We also thank Zheng Zheng for the use of his merger tree code. This work was supported in part by NSF grant NSF-0098515 and NASA grant HST-GO-09838.

#### REFERENCES

- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307
- Bennett, C. L. et al. 2003, ApJS, 148, 1
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
- Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
- Carr, B. J., Bond, J. R., & Arnett, W. D., 1984, ApJ, 277, 445
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
- Damour, A. 2001, Phys. Rev. D, 64, 124013
- Damour, A., & Buonanno, A. 2000, Phys. Rev. D, 62, 064015
- Eisenstein, D. J., & Hu, W. 1999, ApJ, 511, 5
- Efstathiou, G., & Rees, M. J. 1988, MNRAS, 230, 5
- Fan, X. et al. 2003, AJ, 125, 1649
- Favata, M., Hughes, S. A., & Holz, D. E. 2004, ApJ, 607, L5
- Fitchett, M. J. 1983, MNRAS, 203, 1049
- Gould, A., & Rix, H.-W. 2000, ApJ, 532, L39
- Haiman, Z. 2004, ApJ, in press
- Haiman, Z., & Cen, R. 2002, ApJ, 578, 702
- Haiman, Z., & Loeb, A., 2001, ApJ, 552, 459
- Hut, P., & Rees, M. J. 1992, MNRAS, 259, 27p.
- Keeton, C. R., Kuhlen, M., & Haiman, Z. 2004, ApJ, submitted (astro-ph/0405143)
- Kidder, L. E. 1995, Phys. Rev. D, 52, 821
- Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
- Madau, P., & Quataert, E. 2004, ApJ, 606, L17
- Madau, P., Rees, M. J., Volonteri, M., Haardt, F., & Oh, S. P., 2004, ApJ, 604, 484
- Menou, K., & Haiman, Z. 2004, ApJ submitted (astro-ph/0405335)
- Merritt, D., Milosavljević, M., Favata, M., Hughes, S. A., & Holz, D. E. 2004, ApJ, 607, L9
- Ohsuga, K., Mineshige, S., Mori, M., & Umemura, M. 2002, ApJ, 574, 315
- Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
- Ruszkowski, M., & Begelman, M. C. 2003, ApJ, 586, 384
- Salpeter, E. E. 1964, ApJ, 140, 796
- Spergel, D. N. et al. 2003, ApJS, 148, 175
- Volonteri, M., Haardt, F., & Madau, P. 2003, ApJ, 582, 559
- Willott, C. J., McLure, R. J., & Jarvis, M. J. 2003, ApJ, 587, 15
- Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003, ApJ, 592, 645